

CUL MWG Data Scripting Bootcamp

GitHub Repository for the Bootcamp sample data: <https://github.com/cmh2166/CUL-MWG-Workshop>

About this Bootcamp

The Cornell University Library Metadata Working Group is hosting a Data Scripting Bootcamp on **February 20th, from 10:00 AM-4:00 PM**, with hour for lunch. This will be in **Mann Library B30A**, and we request that participants register [here](#).

This Bootcamp is meant to cover a number of (meta)data scripting topics in a short amount of time, with the focus being on learning through practical use. You won't be an expert in these areas by the end of this bootcamp; you won't even be entirely comfortable with many of these concepts at the end of the bootcamp, and this day won't cover everything.

Instead, this bootcamp is meant to be a way for you to jump into using scripting for common metadata needs you might have. This bootcamp aims to give you some preparation with the basic tools for working with metadata applications and scripts that we dive into [further detail in other MWG workshops this Spring](#).

As such, we strongly recommend you register [here](#) and provide use with examples of workflows, use cases, or data you'd like to work with using the skills in this workshop. We will try to work your examples into this day!

Proposed Outcomes for Participants:

- Know basics of data structures with particular relevance for library metadata work at Cornell;
- Know basics of working in a Command Line Interface with bash/shell and examples of working with metadata in this environment;
- Know basics of SQL, with examples and use cases for metadata queries and updates in particular;
- Be able to run simple scripts and programs, understanding how this works with relevance for metadata work;
- Know basics of Git and version control, be able to create a repo, push a change, check version of a dataset, etc.

Bootcamp Details

Pre-requisites

All users will need to create a [GitHub account](#) - and remember your username/password for the day.

We will have the Mann Library B30A lab computers set up all the software needed. If you don't want to use your own laptop and/or prefer Windows, then just create your GitHub Account and you're golden.

If you prefer to work with a Mac (which we would prefer you do if possible), you can either:

- Check out a Mac Laptop from the Mann Library Circulation desk and let the instructor know at least 15 minutes before the workshop starts (so we can run installations for you).
- Bring your own Mac Laptop with the following installed.

If you are bringing your own laptop, this is the software needed to participate. It looks scarier than it is (most of this will already be on your Mac or Linux laptop, and most of this has installers). Note that MySQL is entirely optional. If you have questions during this installation, you can bug [user-9f226](#) via email - cmharlow@gmail.com or CUL-IT Slack ([cmh329](#)). I will respond right away.

software	Mac OSX Installation	Windows (earlier than 10)	Windows 10	Linux
bash shell	Comes pre-installed - look for Terminal (found in /Applications/Utilities) on Mac OS X	Git for Windows (this will provide you with both Git and Bash in the Git Bash program.) 1. Download the Git for Windows installer . 2. Run the installer and follow the steps below: a. Click on "Next". b. Click on "Next". c. Click on "Next". d. Click on "Next". e. Click on "Next". f. Select "Use Git from the Windows Command Prompt" and click on "Next". If you forgot to do this programs that you need for the workshop will not work properly. If this happens rerun the installer and select the appropriate option. g. Click on "Next". Keep "Checkout Windows-style, commit Unix-style line endings" selected. h. Select "Use Windows' default console window" and click on "Next". i. Click on "Next". j. Click on "Finish".	Windows 10 provides users with bash capabilities natively. To set up the Windows 10 (Anniversary Update) Shell, follow the directions here: http://www.windowscentral.com/how-install-bash-shell-command-line-windows-10	Comes pre-installed - The default shell is usually Bash, but if your machine is set up differently you can run it by opening a terminal and typing bash .

Any Modern Web Browser (not Internet Explorer)	Try Firefox or Chrome .	Try Firefox or Chrome .	Try Firefox or Chrome .	Try Firefox or Chrome .
Git (Install & SetUp with your GitHub Account)	Install and set up Git following the directions here . Note - this may require you download XCode Tools first, which can be a lengthy (but not complicated) download.	If you used the Git for Windows installation above for bash shell, then all you need to do is set up Git in that Git for Windows shell. Follow only step 2 forward here .	Git for Windows (this will provide you with both Git.) <ol style="list-style-type: none"> 1. Download the Git for Windows installer. 2. Run the installer and follow the steps below: <ol style="list-style-type: none"> a. Click on "Next". b. Click on "Next". c. Click on "Next". d. Click on "Next". e. Click on "Next". f. Select "Use Git from the Windows Command Prompt" and click on "Next". If you forgot to do this programs that you need for the workshop will not work properly. If this happens rerun the installer and select the appropriate option. g. Click on "Next". Keep "Checkout Windows-style, commit Unix-style line endings" selected. h. Select "Use Windows' default console window" and click on "Next". i. Click on "Next". j. Click on "Finish". <p>After that installation completes, follow only step 2 forward here to set up your installation.</p>	If Git is not already available on your machine you can try to install it via your distro's package manager. For Debian /Ubuntu run <code>sudo apt-get install git</code> and for Fedora run <code>sudo yum install git</code> .
SQLite	SQLite comes pre-installed on Mac OS X.	Download and run <code>sqlite-dll-win32-x86-[some numbers].zip</code> from the SQLite site .	Download and run <code>sqlite-dll-win32-x86-[some numbers].zip</code> from the SQLite site .	SQLite comes pre-installed on Linux.
MySQL (optional)	The recommended way for installing MySQL on OS X is to use the OS X installer package. See Installing MySQL on OS X Using Native Packages on how to download and run the installer package, and how to start the MySQL server afterward. Detailed information regarding installation on OS X can be found in Installing MySQL on OS X .	The recommended way to install MySQL on Microsoft Windows is to use the MySQL Installer; see MySQL Installer Method on how to download and run the MySQL Installer. For a detailed explanation for each step of the installation wizard, see MySQL Installer GUI . If you have chosen to configure MySQL as a Windows service during the installation process, which is the default option (see Windows Service for details), the MySQL server will start automatically after the installation process is completed. Detailed information regarding Windows installation, including alternative installation methods and instructions for troubleshooting, can be found in Installing MySQL on Microsoft Windows .	The recommended way to install MySQL on Microsoft Windows is to use the MySQL Installer; see MySQL Installer Method on how to download and run the MySQL Installer. For a detailed explanation for each step of the installation wizard, see MySQL Installer GUI . If you have chosen to configure MySQL as a Windows service during the installation process, which is the default option (see Windows Service for details), the MySQL server will start automatically after the installation process is completed. Detailed information regarding Windows installation, including alternative installation methods and instructions for troubleshooting, can be found in Installing MySQL on Microsoft Windows .	The easiest way to install MySQL is to use the MySQL repositories: For Yum-based Linux distributions like Oracle Linux, Red Hat Enterprise Linux, and Fedora, follow the instructions in A Quick Guide to Using the MySQL Yum Repository . If your system cannot use the MySQL Yum repository for some reason, follow the instructions in Installing MySQL on Linux Using RPM Packages from Oracle . For APT-based distributions like Debian and Ubuntu, follow the instructions in A Quick Guide to Using the MySQL APT Repository . If your system cannot use the MySQL APT repository for some reason, follow the instructions in Installing MySQL on Linux Using Debian Packages from Oracle . For SUSE Linux Enterprise, follow the instructions in A Quick Guide to Using the MySQL SLES Repository . If your system cannot use the MySQL SUSE repository for some reason, follow the instructions in Installing MySQL on Linux Using RPM Packages from Oracle .

Schedule



Note: These times may shift. If you intend to only go to one session, please let us know by registering [here](#). That way, we can alert you to time shifts if/when they occur for the session of interest to you.

Session	Time	Description
Data Structures & Models	10 - 11 AM	Brief introduction to data structures and models for the metadata worked with in the workshop
*nix Command Line Interface / Shell	11 - 12:30 PM	Overview of the CLI for running simple data scripts. Will introduce some basic bash scripts, as well as how to run existing Python scripts (and begin to debug errors) for analyzing metadata. YOU WON'T LEARN PYTHON, but how to run Python scripts.
<i>Lunch</i>	12:30 - 1:30	
Git for Metadata	1:30 - 2:45 PM	Using Git and GitHub for metadata work run previously.
SQL / MySQL / SQLite	2:45 - 3:45 PM	Built off the previous, learn the basics of SQL - what it does and how to run basic queries or updates on metadata managed or used through out this workshop.
Wrap-up	3:45 - 4 PM	Any remaining questions, proposals for the next workshops, where to access materials or get help, etc.

Gratitude & License

This workshop materials reuses open curricula from the [Library Carpentry project](#). We have updated lessons with specific examples for the Cornell University Library Metadata Working Group examples and data needs. We are not endorsed by Software, Data, or Library Carpentry.

We've also added curriculum materials influenced by similar introduction to scripting workshops at Library Technology conferences, so we would like to give a special thanks to Patrick Hochstenbach (Ghent University Library) & Johann Rolschewski (Berlin State Library) for their [open and shared work](#).

As such, this workshop material is licensed under a [Creative Commons Attribution 4.0 International License](#).